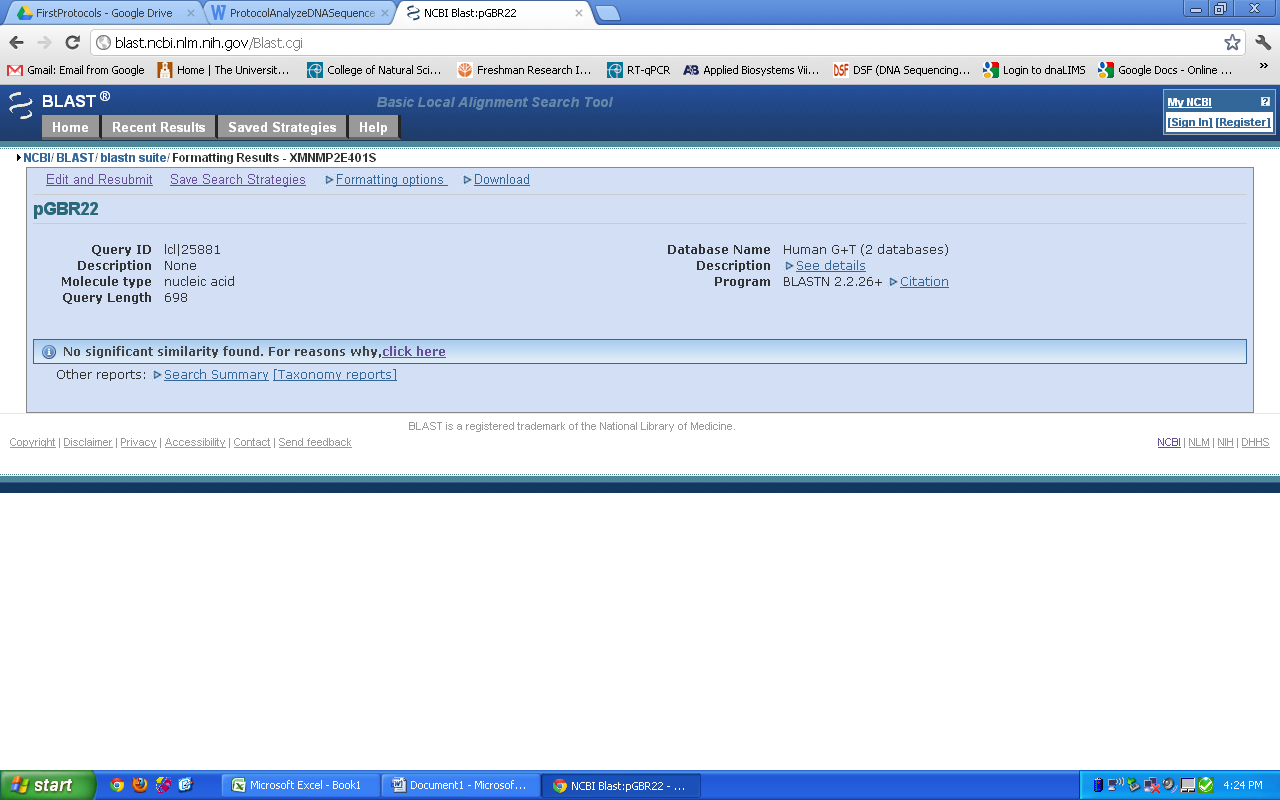
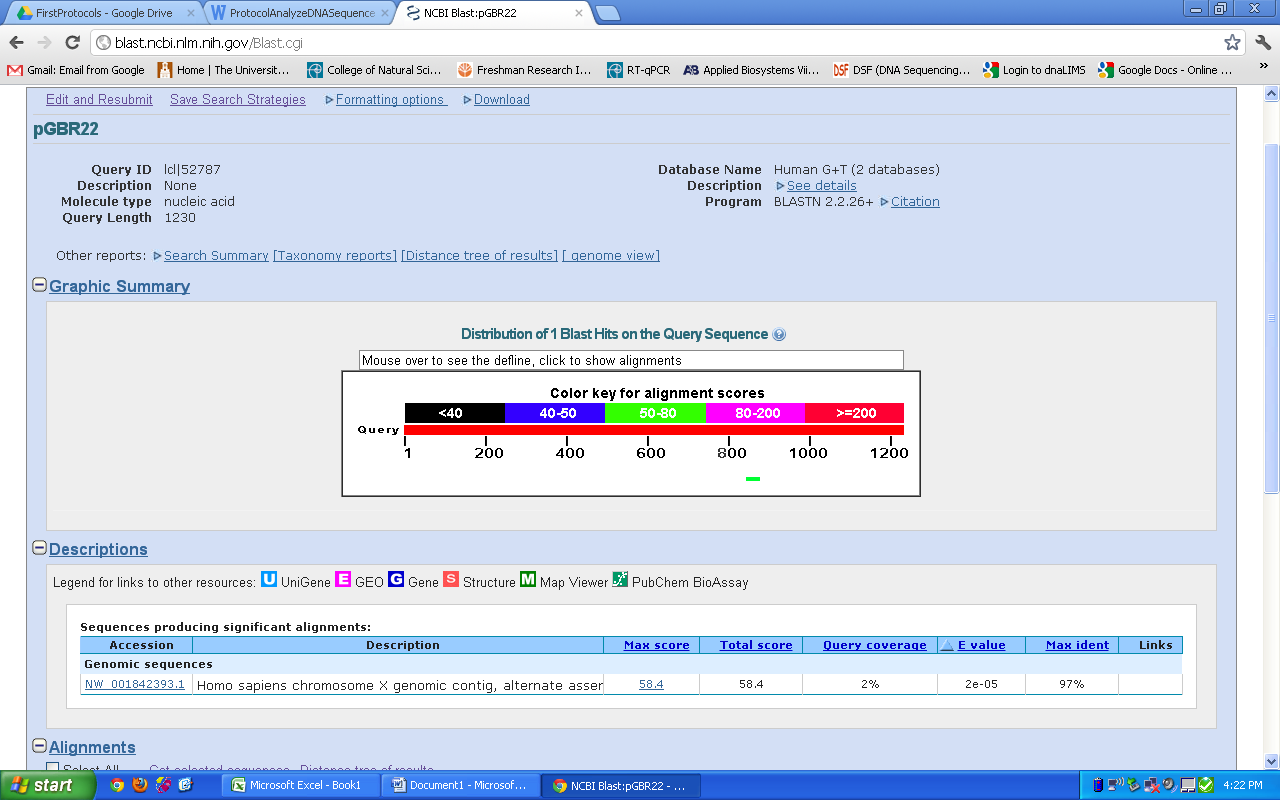
Forward (from DNALIMS):

NNNNNNNNNNNNNNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTNNNNNNNNNNNNNTNGNNNNCTNTNNNNGGNNGCGTTCTTGANGCGACNCNANNNTNTNNNNTTANTTNNNNNGNNTNCCNTGANNGGNNATTGGNGANNNNATGACCTNNNNNGTGTNTGNNTNNNGNNNNNNCANNNNNCACTCNNNNGCNNNCNNANGNNANGGNTNCNNNTGNTNNCNGAAGGGNAGGGGANTNNCANNCCGGNTGTCCCCCNNGGTGCACGNCTNNCATTNNCCTNTNATNNTNTATNCCCNCTNTCNCNATNCGAANNNNTNCAATTCNNCNCNCACCCTGANNACCTCCCTGGGNNNNNANAGCAGCGGTTNCGTGGGGNANGCNCACGGNNNNNNNATTGNNNNCATTNGANNATANNNCTCTGNGNACNANCANNANTGATTCCNNGNNNNNNNGNNNNNGTNTCNGCTNCNANNNNNANTCTNTNGNANGNGCTTNCCTCCCNATGCACCTNNTATGCANNANNNNACNCNGGNNTGNTNNCNCATCACTGANTNTGNNCNTNNNCGACANNNNNGCTGANAGGAANNNNCTTTNNNNCTCTNACTTGGNANGNNGTGGTTACTNCNTCNGTNNTTCNNNNNTNCNGNACNNNNNNNNNANNNCCNANNNNATNNNNNNNNANNNCNA

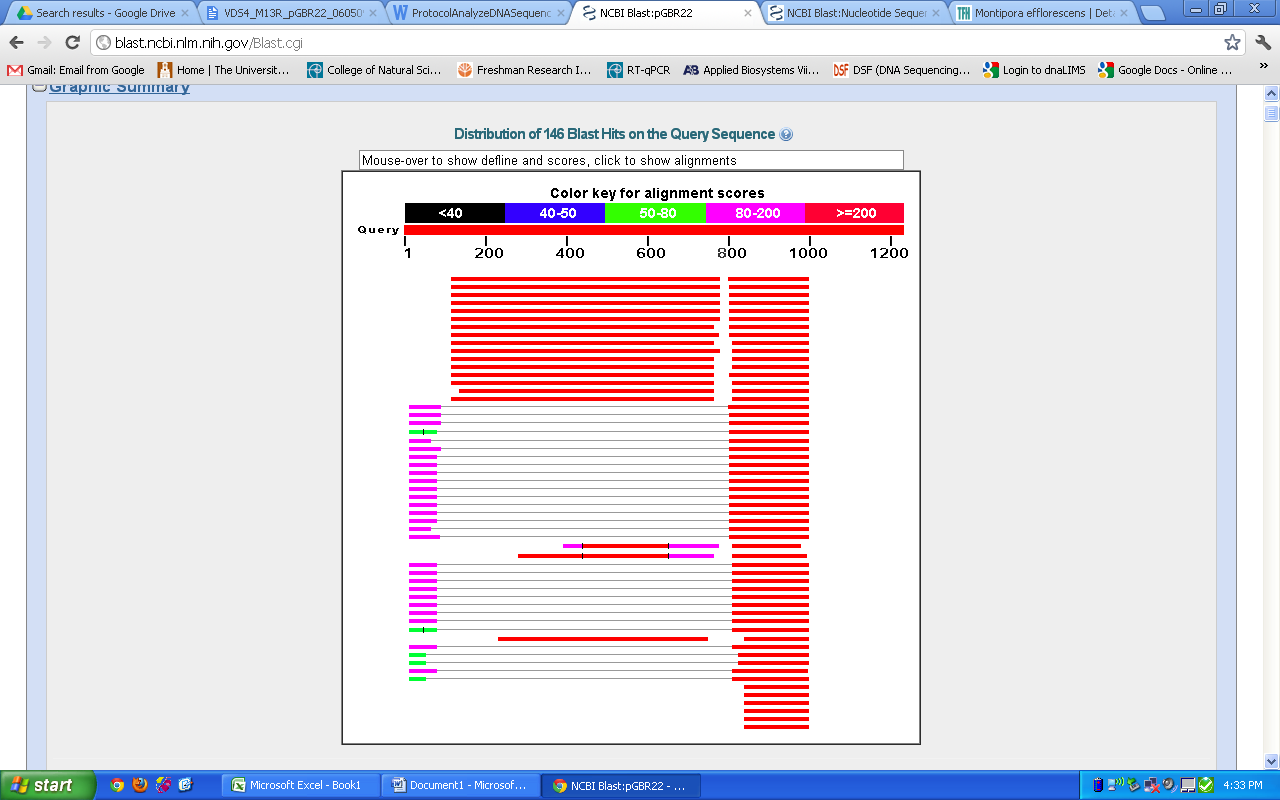
BLAST- Compared to human genome (same result for all genomes):

  
Reverse (from DNALIMS):  
  
NNNNNNNNNNNNNNATAGANTACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCNGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGNGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATANCGANNANGCCCGCANCGATCGCCCNNCCCANAGTNNCNCANCNGNATGNNATGGNCNNNCCNNNANCGNNCNTNNCNNNGCGGGNNNNGNNNNNCNNNNANNNNNACNNTANNNTNNCNGCNNNNNNNNNNNNNNTNNNNNNNNNNNNNNNNNNNNNNNTNNCNNNNCCNNNANNNNNNNNNNGNNNNNNNNNNNNCNANTNNNNNNNNNNNNGNNNNANNNNNNNNNNNNANNNNNNNNNNNNGGCNNNCNNNN

BLAST- Compared to human genome:



pGBR22 does not share much in common when compared to the human genome. Perhaps this is because it is not part of the human genome. To verify this statement, I checked the DNA sequence against all genomes available on the database.

Compared to all genomes:

Montipora efflorescens has the most matches. The DNA sequence I received from DNALIMS was very similar to the genome of Montipora efflorescens.

The alignments were:

Montipora efflorescens GFP-like chromoprotein mRNA, complete

cds

Length=666

Score = 1196 bits (1326), Expect = 0.0

Identities = 663/663 (100%), Gaps = 0/663 (0%)

Strand=Plus/Plus

Query 118 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 177

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 1 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 60

Query 178 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 237

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 61 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 120

Query 238 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 297

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 121 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 180

Query 298 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 357

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 181 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 240

Query 358 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 417

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 241 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 300

Query 418 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 477

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 301 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 360

Query 478 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 537

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 361 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 420

Query 538 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 597

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 421 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 480

Query 598 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 657

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 481 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 540

Query 658 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 717

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 541 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 600

Query 718 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 777

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 601 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 660

Query 778 GGT 780

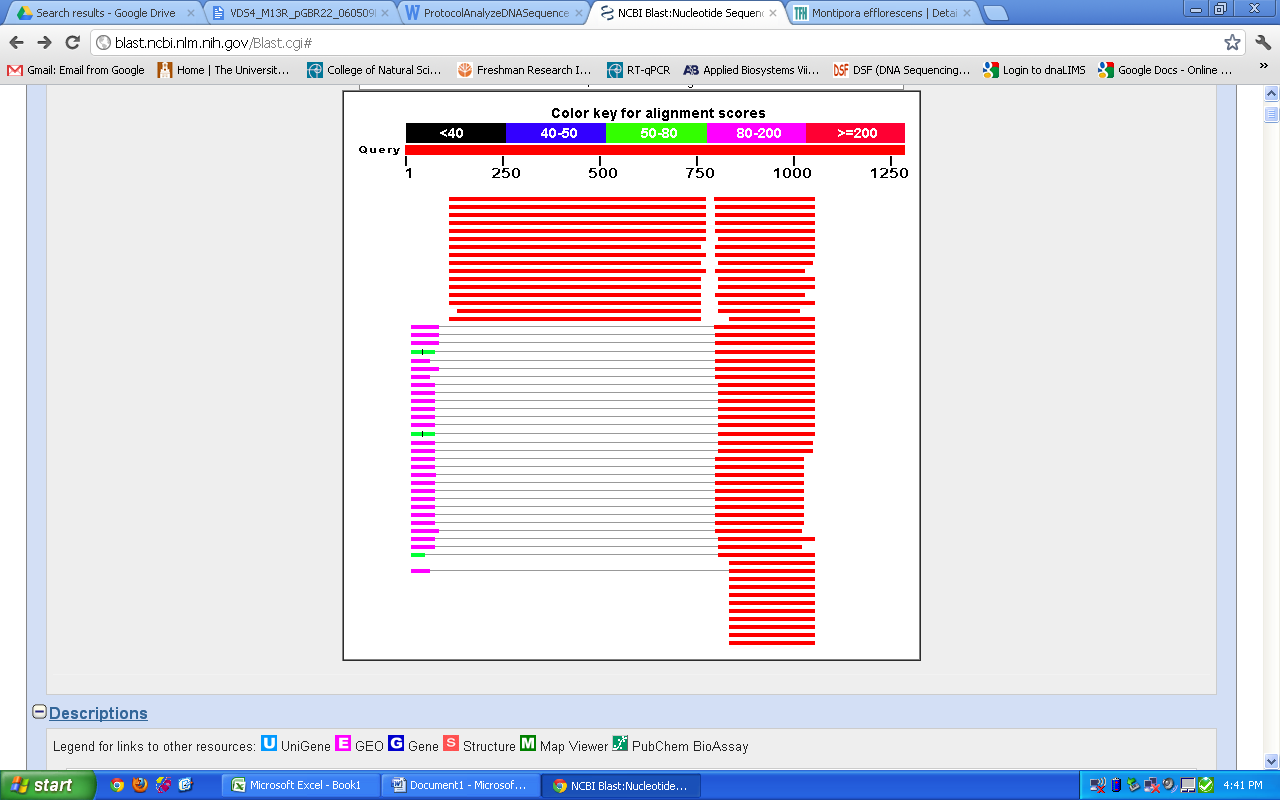
|||

Sbjct 661 GGT 663

Using this sequence (found at VDS4\_M13R\_pGBR22\_060509forBB.txt):

NNNNNNNNNNNTATNGNATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAANAGGCCCGCACCGATCGCCCTTCCCACAGTTGCGCAGCCTGATGGCNATGGACGCNNCCTGNANCGNCGCATTAGCNCGGCNGGNNNGNGNNCNCGCAGCGNGACNCTACANTTNCNGCNNCNANCNNNNNCNTTCNNTTNTNCNNNTNNNNNNNNNNNNNNNTTNCCGNNAGNTNNAANNGGGGNNNNNNNNNATNNGNTNNNNNNNNNNCCAAANTGNNNGNNNNGNNNNNNNNGNNNNNNNNNNNNNNNNNNNNNNNGANCNNNNNNNNNNNNANNNNNNNNNNNCNNNNNAANNNNNCCNNTNNN

The BLAST was as follows:



The alignments were

Montipora efflorescens GFP-like chromoprotein mRNA, complete

cds

Length=666

Score = 1225 bits (663), Expect = 0.0

Identities = 663/663 (100%), Gaps = 0/663 (0%)

Strand=Plus/Plus

Query 116 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 175

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 1 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 60

Query 176 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 235

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 61 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 120

Query 236 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 295

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 121 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 180

Query 296 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 355

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 181 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 240

Query 356 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 415

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 241 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 300

Query 416 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 475

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 301 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 360

Query 476 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 535

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 361 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 420

Query 536 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 595

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 421 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 480

Query 596 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 655

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 481 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 540

Query 656 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 715

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 541 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 600

Query 716 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 775

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

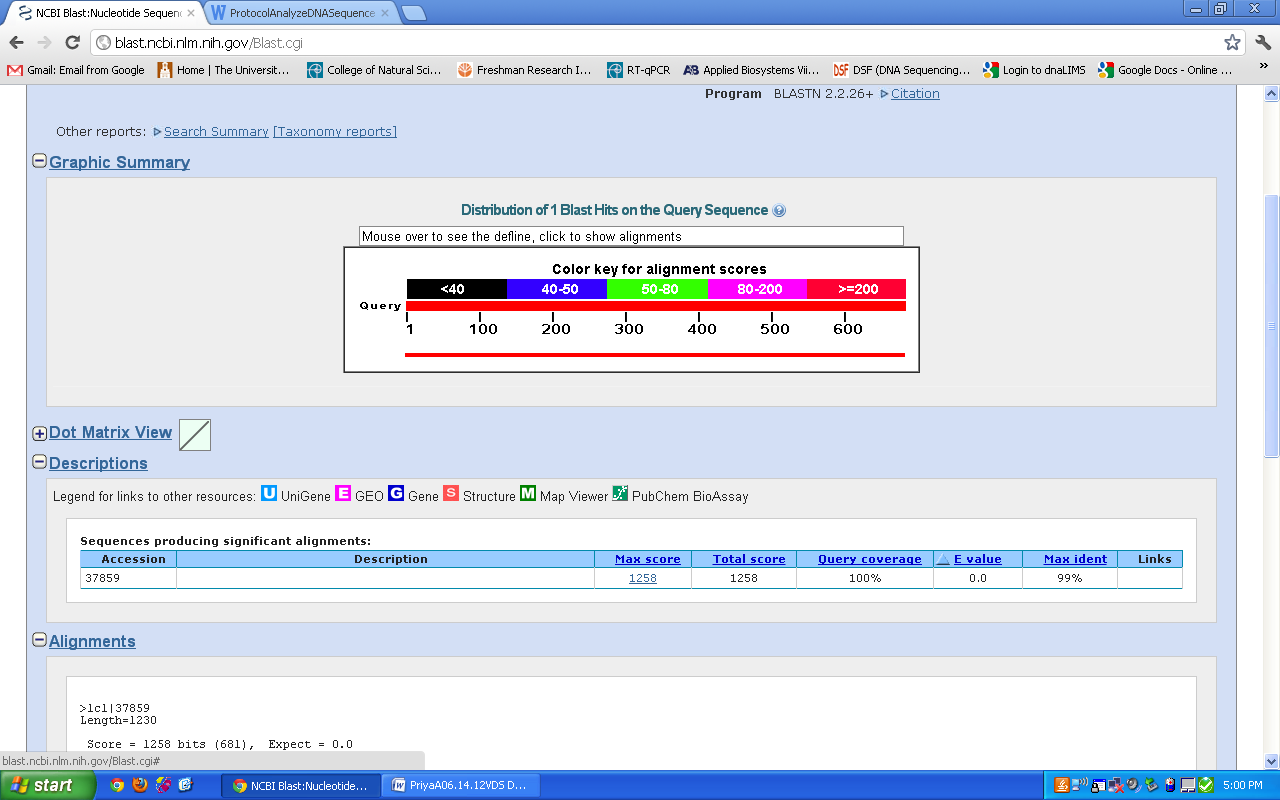
Sbjct 601 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 660

Query 776 GGT 778

|||

Sbjct 661 GGT 663

Known CDS: ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAGTTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAA



Alignments:

>lcl|37859

Length=1230

Score = 1258 bits (681), Expect = 0.0

Identities = 683/684 (99%), Gaps = 0/684 (0%)

Strand=Plus/Plus

Query 1 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 118 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 177

Query 61 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 178 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 237

Query 121 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 238 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 297

Query 181 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 298 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 357

Query 241 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 300

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 358 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 417

Query 301 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 360

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 418 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 477

Query 361 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 420

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 478 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 537

Query 421 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 480

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 538 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 597

Query 481 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAGTTCAAATCTACTTACAAGGCAAAG 540

||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||

Sbjct 598 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 657

Query 541 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 600

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 658 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 717

Query 601 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 660

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 718 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 777

Query 661 GGTCATCACCATCACCATCACTAA 684

||||||||||||||||||||||||

Sbjct 778 GGTCATCACCATCACCATCACTAA 801

<http://www.bioinformatics.org/sms2/>

Translating DNA sequence into protein:

Given sequence (no N’s):

Translate results

>rf 1 Untitled

MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL

SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI

SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK

KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGHHHHHH\*

Translate results

>rf 2 Untitled

\*V\*SLNK\*PTRFICQARSMDTTLRSKAMEKESLTRGSRR\*SSLSPRVDLCHLLGIFYHHC

LNTEAYHSPSTLKTSLIM\*SSHSLRDIHGRGS\*TLKMVQCVLSAMIPASKATVSSTMSKS

LV\*TFLPMDLLCRRRHRAGNPTLSVSLHEMEC\*\*ETTLWL\*SWKEVVTICVSSNLLTRQR

SL\*GCQGITMLTANWM\*PVTTRITHLLSSVKYPLHATLCSVITITIT

Translate results

>rf 3 Untitled

ECDR\*TNDLQGLYVRHGQWTLL\*GRRRWKRKALRGGADGKAHCHQGWTSAICLGYFITTV

SIRKHTIHQVP\*RHP\*LCKAVIP\*GIYMGEDHEL\*RWCSVYCQQ\*FQHPRQLFHLQCQNL

WCELSSQWTCYAEEDTGLGTQH\*ASLCTRWNADRKQLYGSEVGRRWLLFV\*VQIYLQGKE

ACEDARVSLC\*PQTGCNQSQQGLHIC\*AV\*NIHCTPLFARSSPSPSL

Sequence with some N’s:

Translate results

>rf 1 Untitled

STCRRPH\*\*F\*LIEGEIS\*V\*SLNK\*PTRFICQARSMDTTLRSKAMEKESLTRGSRR\*SS

LSPRVDLCHLLGIFYHHCLNTEAYHSPSTLKTSLIM\*SSHSLRDIHGRGS\*TLKMVQCVL

SAMIPASKATVSSTMSKSLV\*TFLPMDLLCRRRHRAGNPTLSVSLHEMEC\*\*ETTLWL\*S

WKEVVTICVNSNLLTRQRSL\*GCQGITMLTANWM\*PVTTRITHLLSSVKYPLHATLCSVI

TITITKIPRPWRPGACDVGPNSPYSESYYNSLAVVLQRRDWENPGVTQLNRLAAHPPFAS

WRNSEXARTDRPS

Translate results

>rf 2 Untitled

RPAGGRTSDFD\*LKEKYHECDR\*TNDLQGLYVRHGQWTLL\*GRRRWKRKALRGGADGKAH

CHQGWTSAICLGYFITTVSIRKHTIHQVP\*RHP\*LCKAVIP\*GIYMGEDHEL\*RWCSVYC

QQ\*FQHPRQLFHLQCQNLWCELSSQWTCYAEEDTGLGTQH\*ASLCTRWNADRKQLYGSEV

GRRWLLFV\*IQIYLQGKEACEDARVSLC\*PQTGCNQSQQGLHIC\*AV\*NIHCTPLFARSS

PSPSLKSRGHGGREHATSGPIRPIVSRITIHWPSFYNVVTGKTLALPNLIALQHIPLSPA

Translate results

>rf 3 Untitled

DLQAAALVILID\*RRNIMSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLT

VTKGGPLPFAWDILSPLSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTV

SNDSSIQGNCFIYNVKISGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKL

EGGGYYLCEFKSTYKAKKPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGHH

HHHH\*NPAAMAAGSMRRRAQFAL\*\*VVLQFTGRRFTTS\*LGKPWRYPT\*SPCSTSPFRQL

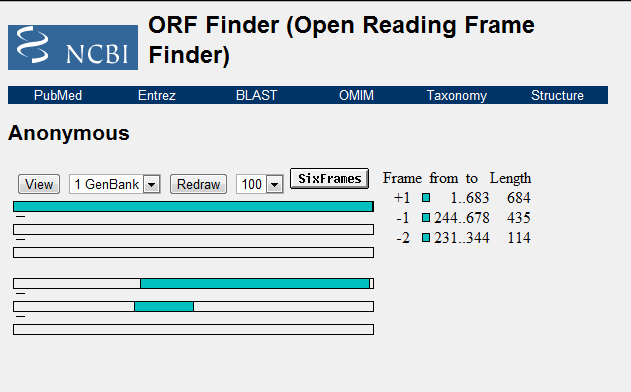
A\*\*RXGPHRSPF

There are three reading frames because codons contain three letters that code for an amino acid. The sequence that is right will be defined as the one that will translate the most amount of DNA (will have the greatest protein output). By this definition, using the given DNA sequence, the correct reading frame is rf1, and using the sequence with some N’s, the correct reading fram is rf2.

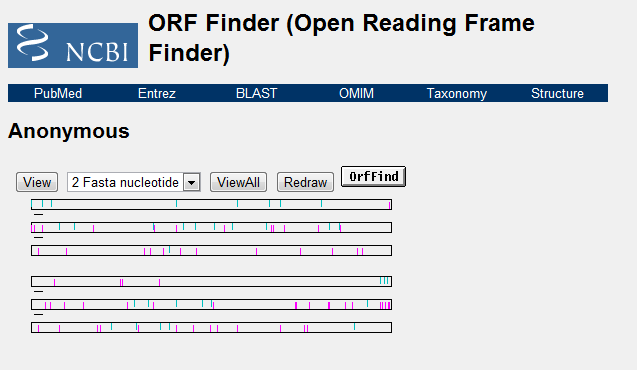
The starting amino acid is Methionine, and is signified by an M. This His6 tag is highlighted in yellow. The His6 tag is the most common form of polyhistidine tags. This His6 tag allows a protein of interest to be latched onto the tag without affecting the activity of the protein. The protein with the His6 tag can then be attached to a protein of interest for further study.

ORF Finder:

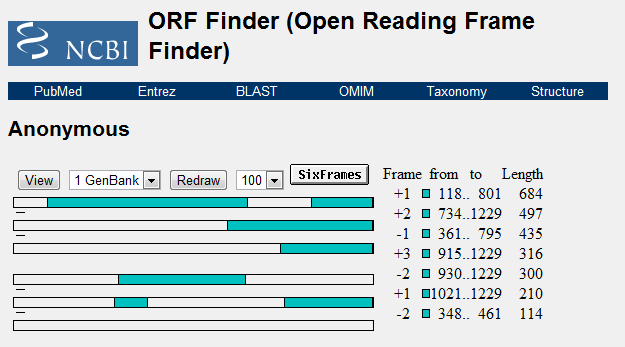
KNOWN:



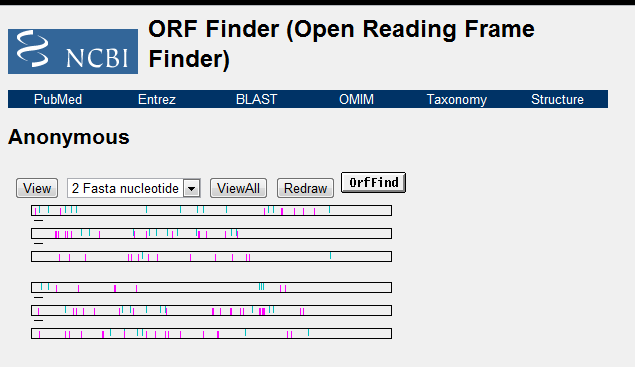
Six Frames:



READ:



6 Frames:



LOCUS tmpseq\_0 1230 bp linear 14-JUN-2012

DEFINITION No definition line found.

ACCESSION tmpseq\_0

VERSION

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

FEATURES Location/Qualifiers

source 1..1230

BASE COUNT 283 a 257 c 252 g 246 t 192 others

ORIGIN

1 nnnnnnnnnn nnnnatagan tactcaagct atgcatccaa cgcgttggga gctctcccat

61 atggtcgacc tgcaggcggc cgcactagtg attttgattg attgaaggag aaatatcatg

121 agtgtgatcg ctaaacaaat gacctacaag gtttatatgt caggcacggt caatggacac

181 tactttgagg tcgaaggcga tggaaaagga aagccttacg agggggagca gacggtaaag

241 ctcactgtca ccaagggtgg acctctgcca tttgcttggg atattttatc accactgtct

301 caatacggaa gcataccatt caccaagtac cctgaagaca tccctgatta tgtaaagcag

361 tcattccctg agggatatac atgggagagg atcatgaact ttgaagatgg tgcagtgtgt

421 actgtcagca atgattccag catccaaggc aactgtttca tctacaatgt caaaatctct

481 ggtgtgaact ttcctcccaa tggacctgtt atgcagaaga agacacaggg ctgggaaccc

541 aacactgagc gtctctttgc acgagatgga atgctgatag gaaacaactt tatggctctg

601 aagttggaag gaggtggtta ctatttgtgt gaattcaaat ctacttacaa ggcaaagaag

661 cctgtgagga tgccagggta tcactatgtt gaccgcaaac tggatgtaac cagtcacaac

721 aaggattaca catttgttga gcagtgtgaa atatccattg cacgccactc tttgctcggt

781 catcaccatc accatcacta aaatcccgcg gccatggcgg ccgggagcat gcgacgtcng

841 gcccaattcg ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg

901 tgactgggaa aaccctggng ttacccaact taatcgcctt gcagcacatc cccctttcgc

961 cagctggcgt aatancgann angcccgcan cgatcgcccn ncccanagtn ncncancngn

1021 atgnnatggn cnnnccnnna ncgnncntnn cnnngcgggn nnngnnnnnc nnnnannnnn

1081 acnntannnt nncngcnnnn nnnnnnnnnn tnnnnnnnnn nnnnnnnnnn nnnntnncnn

1141 nnccnnnann nnnnnnnngn nnnnnnnnnn ncnantnnnn nnnnnnnngn nnnannnnnn

1201 nnnnnnannn nnnnnnnnng gcnnncnnnn

//

ORF is the part of the gene that encodes a protein. The ORF finder helps analyse ORFs from nucleotide sequences. It helps find protein coding regions.

What do the 6 frames represent? The different reading frames.

Compare Protein Translations:

(VDS4\_M13R\_pGBR22\_060509forBB.txt) Reading Frame 3

RRNIMSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLT

VTKGGPLPFAWDILSPLSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTV

SNDSSIQGNCFIYNVKISGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKL

EGGGYYLCEFKSTYKAKKPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGHH

HHHH

(PurpleProteinCodingSequenceforBB.txt) Reading Frame 1

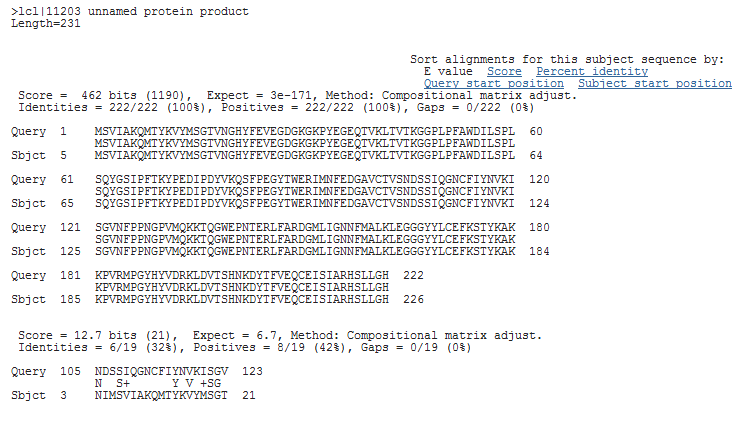
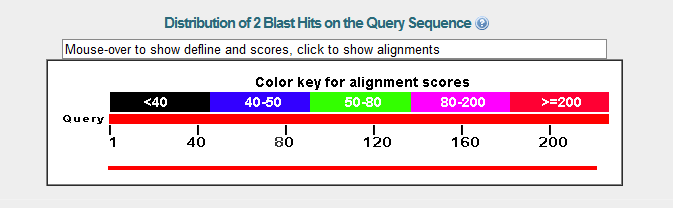
MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL

SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI

SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK

KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGHHHHHH

BLASTP Align Two or more Sequences



Insertion into Vector backbone:

pGEMT sequence (pgemTvectorSequenceforBB.txt)

1 GGGCGAATTG GGCCCGACGT CGCATGCTCC CGGCCGCCAT GGCCGCGGGA  
51 TTAATCACTAGT GCGGCCGCCT GCAGGTCGAC CATATGGGAG AGCTCCCAAC  
101 GCGTTGGATG CATAGCTTGA GTATTCTATA GTGTCACCTA AATAGCTTGG  
151 CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA  
201 ATTCCACACA ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC  
251 CTAATGAGTG AGCTAACTCA CATTAATTGC GTTGCGCTCA CTGCCCGCTT  
301 TCCAGTCGGG AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC  
351 GCGGGGAGAG GCGGTTTGCG TATTGGGCGC TCTTCCGCTT CCTCGCTCAC  
401 TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT  
451 CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG  
501 AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC  
551 GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCCTGACGAG CATCACAAAA  
601 ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC  
651 CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT  
701 GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC  
751 TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC  
801 TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC  
851 CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT  
901 CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA  
951 GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG  
1001 AAGAACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA  
1051 AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT  
1101 GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA  
1151 AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA  
1201 ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC  
1251 TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA  
1301 TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA  
1351 TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC  
1401 GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC  
1451 AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA  
1501 ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC  
1551 GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC  
1601 GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG  
1651 TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA  
1701 TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC  
1751 CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC  
1801 TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA  
1851 AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA  
1901 GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA CGGGATAATA  
1951 CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT  
2001 TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT  
2051 GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA  
2101 GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA  
2151 ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTCAATA  
2201 TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG  
2251 AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA  
2301 AAAGTGCCAC CTGATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA  
2351 AATACCGCAT CAGGAAATTG TAAGCGTTAA TATTTTGTTA AAATTCGCGT  
2401 TAAATTTTTG TTAAATCAGC TCATTTTTTA ACCAATAGGC CGAAATCGGC  
2451 AAAATCCCTT ATAAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTTGT  
2501 TCCAGTTTGG AACAAGAGTC CACTATTAAA GAACGTGGAC TCCAACGTCA  
2551 AAGGGCGAAA AACCGTCTAT CAGGGCGATG GCCCACTACG TGAACCATCA  
2601 CCCTAATCAA GTTTTTTGGG GTCGAGGTGC CGTAAAGCAC TAAATCGGAA  
2651 CCCTAAAGGG AGCCCCCGAT TTAGAGCTTG ACGGGGAAAG CCGGCGAACG  
2701 TGGCGAGAAA GGAAGGGAAG AAAGCGAAAG GAGCGGGCGC TAGGGCGCTG  
2751 GCAAGTGTAG CGGTCACGCT GCGCGTAACC ACCACACCCG CCGCGCTTAA  
2801 TGCGCCGCTA CAGGGCGCGT CCATTCGCCA TTCAGGCTGC GCAACTGTTG  
2851 GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT ATTACGCCAG CTGGCGAAAG  
2901 GGGGATGTGC TGCAAGGCGA TTAAGTTGGG TAACGCCAGG GTTTTCCCAG  
2951 TCACGACGTT GTAAAACGAC GGCCAGTGAA TTGTAATACG ACTCACTATA

Filter DNA results

>filtered DNA sequence consisting of 3002 bases.

GGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCCATGGCCGCGGGATTAATCACTAGTGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA

AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG

GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATA

Reverse Complement results

>Untitled reverse complement

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCC

M13Reverse Read

NNNNNNNNNNNNNNATAGANTACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCNGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGNGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATANCGANNANGCCCGCANCGATCGCCCNNCCCANAGTNNCNCANCNGNATGNNATGGNCNNNCCNNNANCGNNCNTNNCNNNGCGGGNNNNGNNNNNCNNNNANNNNNACNNTANNNTNNCNGCNNNNNNNNNNNNNNTNNNNNNNNNNNNNNNNNNNNNNNTNNCNNNNCCNNNANNNNNNNNNNGNNNNNNNNNNNNCNANTNNNNNNNNNNNNGNNNNANNNNNNNNNNNNANNNNNNNNNNNNGGCNNNCNNNN

Inserting M13Reverse Read into pGemT reverse complement:

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACA**CAGGAAACAGCTATGAC**ATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCC

The M13 Reverse (-27) primer site

Underlined and bold **CAGGAAACAGCTATGAC**

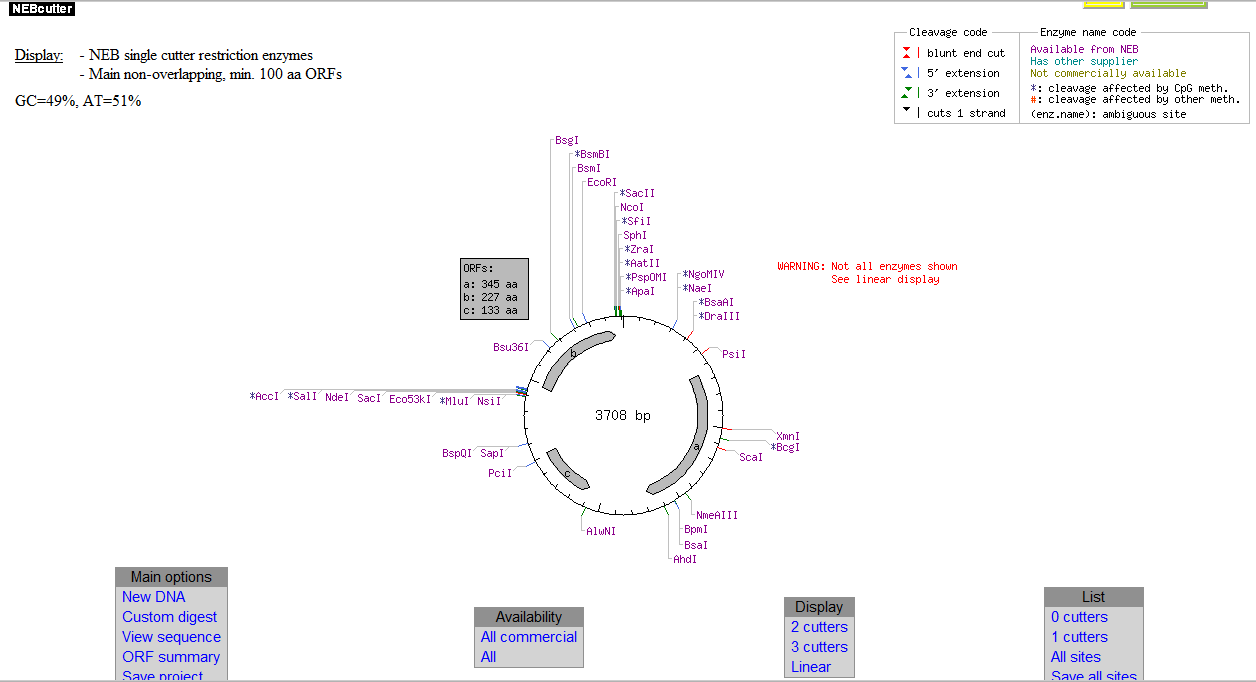
Slicing & Dicing: Now that you have the complete sequence for your pGBR22, you can

determine which Restriction Enzymes might cut it.

Task: cut the plasmid into different size pieces

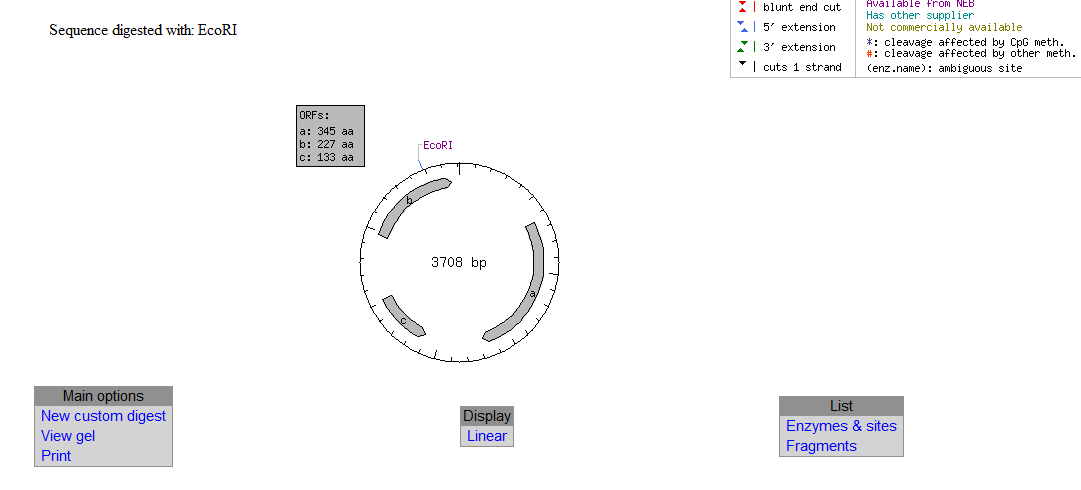
The reason for this is so that they will be distinguishable when run on a gel.

NEB Cutter on the NEB website

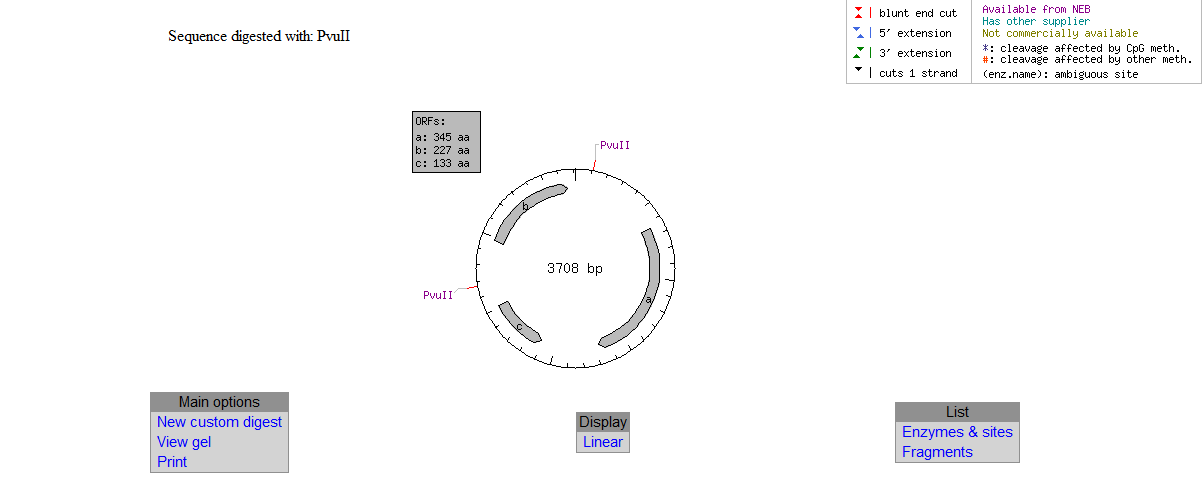


The grey regions correspond to ORFs.

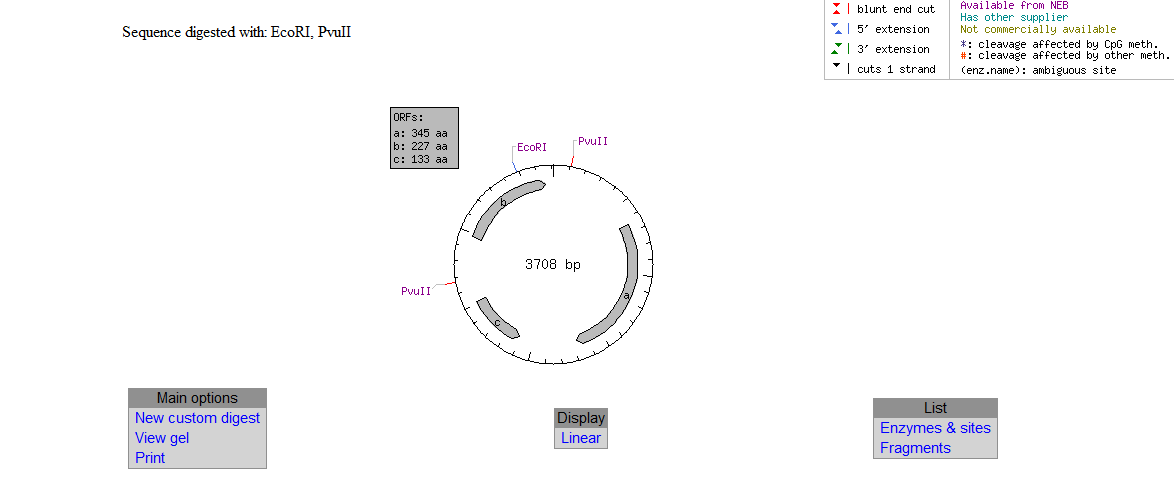
1. cut with EcoRI



2. with PvuII



3. with EcoRI+PvuII at the same time



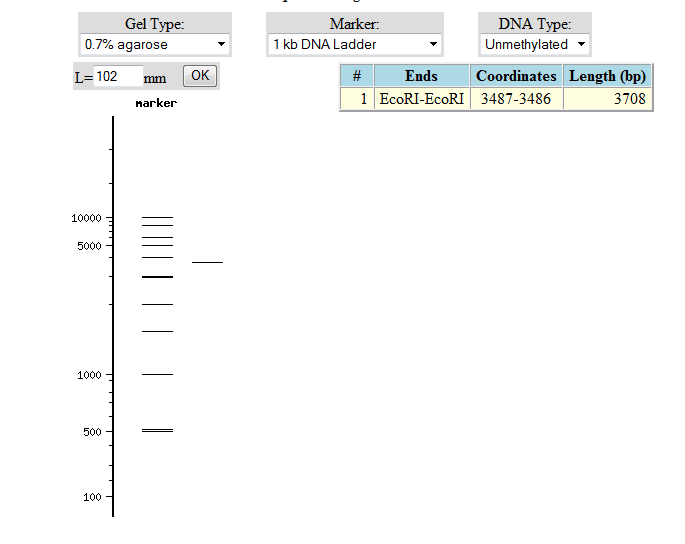
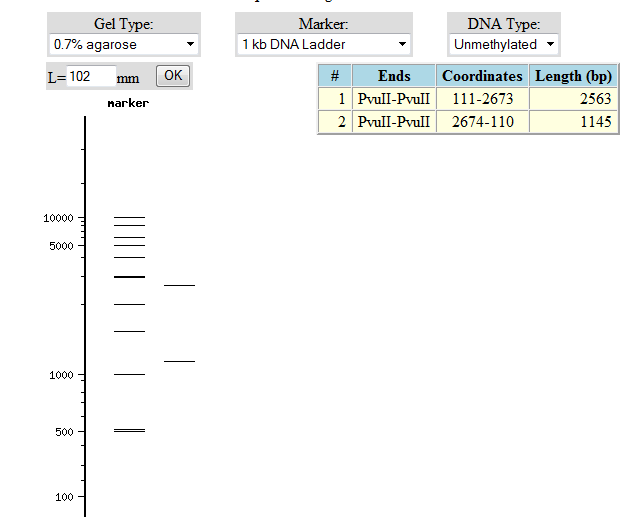
What buffers do these enzymes take? EcoRI takes NEBuffer1, and PvuII takes NEBuffer 2.

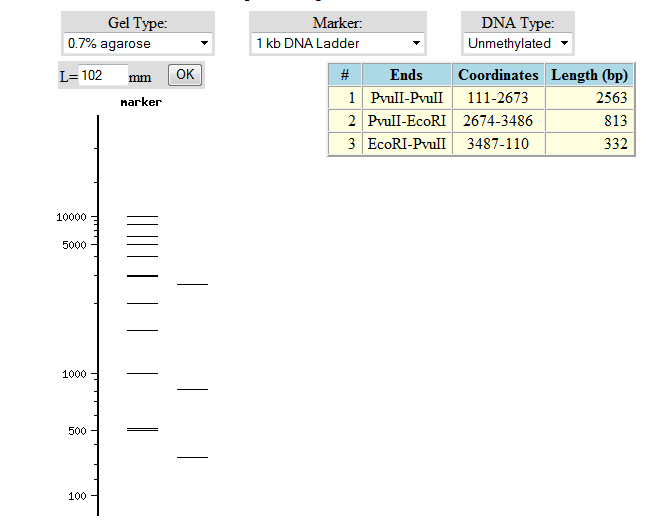
Could they be run together in the same reaction? Yes, because EcoRI is effective in both buffers 1 and 2.

See if the enzymes are ones with a star (\*) - these are methylation sensitive

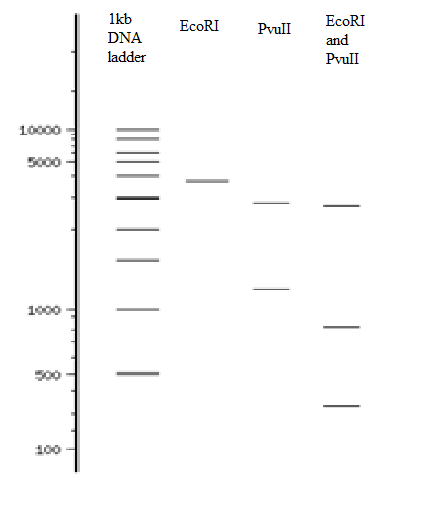
(CpG methylation is ok because it is for mammalian cells, but Dam or Dcm are less desirable because they are in bacterial cells)

- look up Methylation on the NEB cutter site



Virtual Gel



Questions:

How would you determine what your target sequence was if it is bigger than the 1000bp limit of the DNA Sequencing machines? (That is how we got the sequence in the beginning of this exercise)

The primers used in the DNA sequencing at dnaLIMS were M13For-20 and M13Rev-24. These are not the only primers that can be used to sequence DNA. Specific primers can be created according to the amount of DNA to be read. In addition, the DNA can be run multiple times to read more than the 1000bp limit.